8-24-2017

Combined Universal Kriging and Kalman Filter Techniques to Improve Wind Speed Prediction for Northeastern U.S.

Alexander Samalot
alexander.samalot@uconn.edu

Combined Universal Kriging and Kalman Filter Techniques to Improve Wind Speed Prediction for Northeastern U.S.

Alex Samalot

B.S. University of Connecticut, 2014

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Masters of Science

At the

University of Connecticut

2017

Approval Page

Master's of Science Thesis

Combined Universal Kriging and Kalman Filter Techniques to Improve Wind Speed Prediction for Northeastern U.S.

Presented by

Alexander Samalot, B.S.

Major Advisor_____

Marina Astitha

Associate Advisor_____

Emmanouil Anagnostou

Associate Advisor_____

Amvrossios Bagtzoglou

University of Connecticut

2017

Acknowledgements

First and foremost I would like to express gratitude to my major advisor Dr. Astitha for the guidance, dedication and support. Many additional thanks go out to everyone who has helped me along the way, including but not limited to Dr. Anagnostou, Dr. Bagtzoglou, and the entire Atmospheric and Air Quality Modeling Group

Table of Contents

1. Introduction ……1

2. Data ……3

2.1 Observations……3

2.2 NWP Model Outputs ……6

3. Methodology ……9

3.1 Spatial Interpolation (Universal Kriging) ……..9

3.2 Systematic Error Removal (Kalman Filter) ……14

3.3 Combining Universal Kriging (UK) and Kalman Filter (KF) ……17

3.3.1 The First Methodology ……17

3.3.2 The Second Methodology ……18

3.4 Data Processing and Accuracy……...20

4. Results……..…24

4.1 UK Wind Speed Interpolation (Method 1)…….……...24

4.2 RK Residual Interpolation (Method 2)…….……..…29

4.3 Kalman Filter Performance …….……..33

4.4 Overall Model Performance ……...40

5. Conclusions ……50

6. References ……..……53

Abstract

The scope of this study is to identify and improve wind speed prediction errors for storms that have impacted the Northeastern United States during 2003-2014. Accurate wind speed prediction under storm occurrences is significant to identify and assess impacts to the environment and critical infrastructure. Post-processing of a numerical weather prediction model (Weather Research and Forecasting-WRF) was used in the form of Universal Kriging for spatial interpolation and Kalman Filter for bias reduction. Two strategies for using the Kalman Filter in combination with Universal Kriging are investigated and assessed. Universal Kriging of Kalman Filter corrections reduced all error statistics of the WRF model surface wind speed outputs used in this study. The spatial and seasonal variability of wind speed error reduction are also discussed as well as suggestions for future research directions in this topic.

1. Introduction

There has been a long history of using statistical post-processing techniques to calibrate Numerical Weather Prediction (NWP) model output (Hamill et al., 2000; Galanis et al., 2006; Louka et al., 2008; Roux et al., 2010; Delle Monache et al., 2011 among others). Post-processing of NWP improves forecast through removing systematic errors (Hamill et al., 2000). This study aims to identify and improve surface wind speed prediction error for weather storms that have impacted the Northeastern United States (NEUS) from 2003 to 2014. Weather predictions from the Weather Research and Forecasting Model (WRF) (Skamarock et al., 2008) are used in this study.

The WRF model has presented a high surface wind speed bias over land since the early versions of the model (Cheng and Steenburgh, 2005). Despite constant improvement, numerical prediction of atmospheric processes is affected by imperfect initial and boundary conditions, oversimplified terrain, numerical approximations and simplifications of other chemical and physical processes (Wyszogrodzki et al., 2013). WRF model accuracy has been the focus of several studies due to the need for high resolution wind speed forecasts in fields such as power production, air quality, and fire prediction (Cheng and Steenburgh 2005; Roux et al., 2010; Wyszogrodzki et al., 2013; Jiménez and Dudhia 2012).

Surface wind speed evaluations generally result in a positive bias, but are subject to a number of factors (Cheng and Steenburgh 2005; Roux et al., 2010; Wyszogrodzki et al., 2013; Jiménez and Dudhia 2012). Jiménez and Dudhia (2012) demonstrated that the overall bias was greatly influenced by station geography. In their study of the Iberian Peninsula there was a 1.06 m/s bias that was compensated by

a bias of -2.93 m/s at hills and mountains. Overall bias was a consequence of the number of samples taken within certain geographic properties such as elevation (Jiménez and Dudhia 2012). WRF wind speed bias is also affected by the diurnal cycle (Cheng and Steenburgh 2005; Wyszogrodzki et al., 2013; Jiménez and Dudhia 2012) and season (Wyszogrodzki et al., 2013). In addition to bias, WRF model accuracy is reduced by random forecast error caused by the chaotic nature of atmospheric behavior. This cannot be reduced through post-processing techniques used in this study.

Recent studies attempt to correct wind speed biases by accounting for the unresolved topographic features (Mass and Ovens 2011), and adding a new surface sink term in the WRF momentum equation (Jiménez and Dudhia 2012). Post-processing techniques have been tested as well, such as the Analog Kalman Filter (Delle Monache et al., 2011). The scope of this thesis is to examine the temporal and spatial error characteristics of WRF surface wind speed forecasts of 107 storms that impacted NEUS. Bias is measured using in situ wind speed observations from the National Weather Service (NWS) and the Meteorological Assimilation Data Ingest system (MADIS). MADIS is a meteorological observational database of National Oceanic and Atmospheric Administration (NOAA) and quality checked non-NOAA sources (madis.noaa.gov). The experimental design includes two methods of bias reduction for the forecast domain using Universal Kriging and the Kalman Filter. The two strategies for reducing bias across the domain are different due to the use of Universal Kriging to interpolate wind speeds (Luo et al. 2008; Zlatev et al., 2010; Cellura et al., 2008) or residuals (Cellura et al., 2008).

Kriging was selected over deterministic methods (trend surface analysis, inverse distance weighting, local polynomial and thin plate spline) due to the study of seven interpolation methods by Luo et al., (2008). The study found kriging methods produced the most accurate results of daily mean wind speeds, and has been the basis for the use of kriging in several studies (Joyner et al., 2015; Sliz-Szkliniarz and Vogt 2011; Zlatev et al. 2010). The Kalman Filter was selected because it does not require extensive database management (McCollor and Stull 2008) and yields significant model improvement (Louka et al 2008; Delle Monache et al., 2011; Galanis et al., 2006).

Universal Kriging was specifically chosen due to the success in combined models (Cellura et al., 2008; Qian et al., 2014). The study performed by Qian et al. (2014) combines spatial correlation (UK) with temporal information (Bayesian Dynamic Model) to allow a wind farm to make short term wind speed predictions. The study performed by Cellura et al., (2008) uses a Neural Network to create some of the terms used in the Universal Kriging predictor. In simple terms Cellura et al., (2008) kriged the residuals of the Neural Network. Both studies provided very promising results. This thesis compares two methods of a combined post-processing of UK and the Kalman Filter.

2. Data

2.1 Observations

In order to utilize observations of wind speed for the post-processing of NWP outputs, 1 hour 10-m wind speed data from two federal databases (MADIS and

METAR) were collected. The Meteorological Terminal Aviation Routine Weather Report (METAR) is the Aviation Routine Weather Report used by the Federal Aviation Administration (FAA) (faa.gov/regulations_policies). Available stations within the Northeast U.S. domain used in this study are shown in Figure 1a. Wind speed is more variable over shorter distances than other meteorological variables such as temperature and relative humidity (Luo et al., 2008). Due to the size of the forecast domain (Fig. 2) and the highly varying nature of wind speed, an additional database is used to supplement the METAR data. The Meteorological Assimilation Data Ingest System (MADIS) ingests data from NOAA data sources and non-NOAA providers, decodes the data then encodes all of the observational data in common format (madis.noaa.gov). Total MADIS station availability over the domain is shown below (Fig. 1b).

(a) (b)

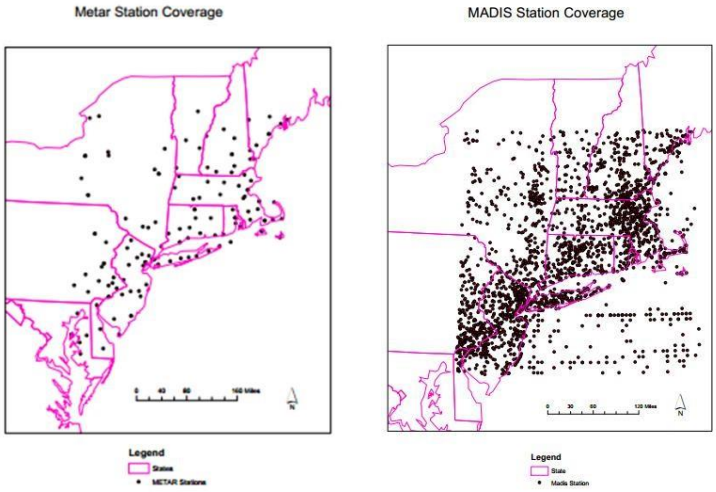


Figure 1: Spatial distribution of station data availability for (a) METAR and (b) MADIS.

Differences in quality control of the data exist between databases. METAR wind speed data meets strict FAA standards. Although MADIS quality checks are not as stringent, MADIS leverages partnerships with international agencies, federal state and local agencies , universities, volunteer networks and the private sector to integrate and quality check observations from their stations with those of NOAA (madis.noaa.gov). Furthermore, the use of MADIS data for WRF wind speed bias behavior and reduction has been seen in Roux et al., (2010). Other recent publications using WRF outputs and MADIS observations include (Lorenzana et al., 2015; Heath et al., 2016)

Figure 1 represents the total stations which contribute to the observed wind speeds in this study. Not all stations in Figure 1 had a full record for all storms without missing data. Maximum data availability is favorable for Kriging and temporal consistency is required for the Kalman Filter and accuracy measurements. In order to maximize availability for these calculations, station data was analyzed for gaps of 5 hours. If the gap was larger than 5 hours from 00 to 23 UTC the station was not Kalman Filtered or cross-validated. If the gap in wind speed was less than 5 hours the data was filled in with an average of the hour before and after as recommended by the EPA for use in regulatory air quality models (Atkinson and Lee 92). Furthermore Figure 1b includes oceanic data, which was removed from interpolation and cross validation due to limited oceanic station coverage.

2.2 NWP Model Outputs

Wind speed NWP forecasts were obtained from the Weather Research and Forecasting model (WRF). WRF is a state of the science mesoscale modeling system. It was designed to create a common software architecture for new and existing models (Skamarock et al., 2008). It is suitable for use in a wide range of applications and scales, ranging from meters to thousands of kilometers (Michalakes, et al. 2005). Organizations involved with WRF include the National Center for Atmospheric Research (NCAR), Air Force Weather Agency (AFWA), National Centers for Environmental Prediction (NCEP), National Oceanic and Atmospheric Administration (NOAA), additional government agencies and various Universities (Cheng and Steenburgh 2005).

The WRF model was applied to the Northeastern U.S to simulate past extreme weather events associated with high wind speeds, intense precipitation and/or snowfall as well as tropical storm systems that affected the area. Below is the map of the three gridded domains of WRF outputs used in this study (Fig. 2). These outputs come from WRF configured with NOAH for land surface scheme (Tewari et al. 2004), Yonsie scheme for PBL (Hong et al. 2006) and Grell 3D scheme for convective parameterization (Grell and Devenyi 2002). The simulations were conducted for another research project by another graduate student (Dr. Maria Frediani, advisor: Prof. E.N. Anagnostou) and the surface wind speed results for the inner domain (2 x 2 km grid spacing) are used in this study.

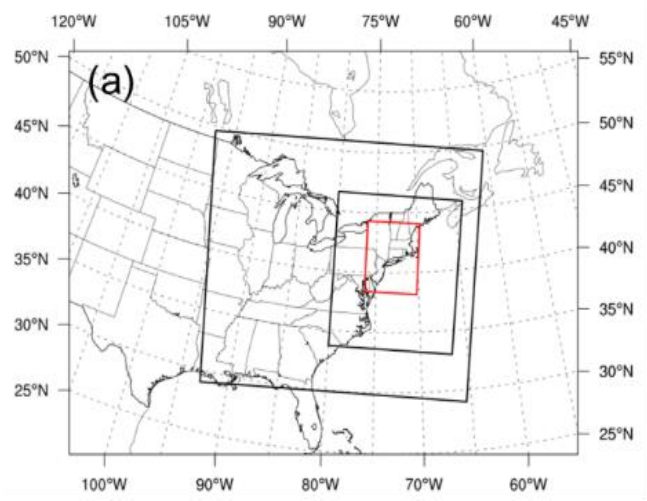


Figure 2: Domain configuration of WRF simulations.

The forecast area is a region of NEUS including a large coastal area, parts of the Appellation Mountains and a wide variety of land types. WRF simulates extreme weather events at various dates from 2003 to 2014 on an event base, with a total of 107 storms used in this study. The database of WRF predictions provides a valuable resource for evaluating model bias behavior. Analysis of forecast errors can help model developers identify the limitations and make improvements to the model (Wyszogrodzki et al., 2013).

WRF surface wind speed evaluations generally result in a positive bias (Cheng and Steenburgh 2005; Roux et al., 2010; Wyszogrodzki et al., 2013; Jiménez and Dudhia 2012). Jiménez and Dudhia showed that the station geography influenced the overall bias. In their study of the Iberian Peninsula there was a 1.06 m/s bias that was compensated by a bias of -2.93 m/s at hills and mountains. Uneven sampling of the wind by the large number of stations located in plains and valleys outweighed mountainous stations yielding a positive bias. Similar geographic differences in WRF

bias were found in a study over the contiguous United States by Wyszogrodzki et al., (2013).

The most notable cause in wind bias differences is terrain elevation, which was designated in many studies (Jiménez and Dudhia 2012; Cheng and Steenburg 2005; Wyszogrodzki et al., 2013). Due to the effect of elevation and other geographic similarities, stations in the western portion of the United States (Rocky Mountains) had large negative biases as low as -3 m/s depending on cycle (Wyszogrodzki et al., 2013). Positive wind speed biases of up to 2 m/s were recorded along Eastern coastal areas.

Temporal factors are another major contributor to bias. Diurnal cycle in wind speed bias was investigated in a number of studies (Cheng and Steenburgh 2005; Roux et al., 2010; Wyszogrodzki et al., 2013; Jiménez and Dudhia 2012). Wyszogrodzki et al., (2013) stated surface wind speed is strongly positively biased at night during the entire study period with averaged biases ranging from 0.5-1 m/s with peaks up to 2 m/s at times. General diurnal behavior had stronger over prediction at night compared to the rest of the day. Furthermore this behavior was affected by seasonality. For spring and summer months (Wyszogrodzki et al., 2013) identified strong negative biases in the afternoon with peak -1.5 m/s which was attributed to limitations in afternoon convection.

3. Methodology

3.1 Spatial Interpolation (Universal Kriging)

There have been a number of studies of wind speed interpolation due to the rise in popularity of wind power (Sliz-Szkliniarz and Vogt 2011; Cellura et al., 2008) and the need for wind speed as input to other models such as pathogens, and insurance/reinsurance (Joyner et al., 2015; Zlatev et al., 2010; Luo et al., 2008). Interpolation of in situ wind speed provides a continuous wind speed surface which incorporates the differences between stations (Akkala et al., 2010; Luo et al., 2008; Joyner et al., 2015). Accuracy of a meteorological spatial interpolation is affected by properties of the meteorological parameter, sensor density and accuracy of the interpolation method chosen (Zlatev et al., 2010). Many methodologies for interpolation exist, deterministic and geostatistical alike.

Limited wind speed interpolation studies have been performed on a wide variety of methods (deterministic and geostatistical). Luo et al., (2008) found that geostatistical methods (such as Kriging) consistently outperformed deterministic methods (trend surface analysis, inverse distance weighting, local polynomial, and thin plate spline). However, there is no general consensus on the best methodology for spatial interpolation of wind speed on a mesoscale heterogeneous area (Joyner et al., 2015). The majority of the literature of wind speed interpolation uses some form of Kriging (Luo et al. 2008; Sliz-Szkliniarz and Vogt 2011; Joyner et al., 2015; Cellura et al., 2008; Zlatev et al. 2009; Zlatev et al. 2010; Qian et al., 2014).

Kriging relies on the basic principal of spatial correlation (Luo et al., 2008). The spatial correlation can be represented by the semivariance function (Eq. 1):

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(s_i) - Z(s_i + h)]^2 \qquad (1)$$

The semivariance function value $\gamma(h)$ at distance h, is a function of the difference between all pair combinations N(h) of an observed wind speed $Z(s_i)$ and other wind speeds at distance h, $Z(s_i + h)$. The semivariance can be directional, accounting for direction dependent variability (Luo et al., 2008). Furthermore, the semivariance function can describe variation that exists within a distance smaller than the distance within sample points. This occurs when the y intercept of the semivariance function is greater than zero, referred to as the nugget effect (Luo et al. 2008; Joyner et al., 2015).

As an interpolator, Kriging has many desirable features. A relatively low amount of data is required, Kriging has the ability to describe anisotropic behavior and Kriging produces an error estimate for every location on the surface. Kriging comes in a variety of forms, but the hypothesis at the basis of the Kriging predictors only involve the first two moments of the spatial field (Cellura et al., 2008). This decreases the sample size required to meet the requirements of the model. Additionally anisotropic (directional) behavior has been a noted property in the wind speed interpolation in Joyner et al., (2015). In addition to an interpolated surface, Kriging also produces estimator error for each location on the surface (Joyner et al., 2015; Cellura et al., 2008).This allows for improved assessment of the field accuracy, identification of troublesome areas and incorporation of Kriging uncertainty into other models.

Even within Kriging, patches of inconsistent wind speed can produce areas of the surface which do not represent observed behavior (Joyner et al., 2015). Kriging is not a perfect estimate for wind speed, and the best way to interpolate a wind speed field of this size is uncertain. Wind speed is more variable over shorter distances than other meteorological variables such as temperature and relative humidity (Luo et al., 2008). Station microclimate may be present when examining macro climatological wind patterns (Klawa and Ulbrich 2003). In high wind events, wind flows relatively uniformly across flat and smooth terrain (Tieleman 1992), but at coastal zones (or other rapid terrain changes) and complex topography, wind speed varies based on the local variability (Tieleman, 1992; Joyner et al., 2015).

Due to the highly variable nature of wind speed, several studies (Luo et al 2008; Joyner et al., 2015; Sliz-Szkliniarz and Vogt 2011) have attempted to find covariate information from variables in addition to distance (coKriging). This has been met with mixed results. In a study of wind speed in Poland, Sliz-Szkliniarz and Vogt (2011) found only a 0.1 correlation to elevation. This was attributed to limited data for the full range of elevation changes in the landscape. Joyner et al., (2015) compared every individual and combination of the direction the slope of the terrain (aspect), land cover and elevation. Each storm in Joyner et al., (2015) responded to different covariates and combinations. coKriging was not used in this study due to the limited data availability as well as the spatial and temporal scope of this investigation.

In similar studies with comparisons between coKriging and UK (Luo et al., 2008; Zlatev et al. 2010) there was no difference in mean error magnitude and a 0.3 m/s difference in RMSE. The most notable drawback to UK was the limitation of the

range of wind speed (Luo et al. 2008; Zlatev et al., 2010). UK calculates trend information in addition to spatial correlation. The trend information can decrease the ability to capture local behavior when the trends are not of benefit. The UK was selected for this thesis due to the success in combined models (Cellura et al., 2008; Qian et al., 2014). Cellura et al., (2008) found UK had lower accompanying prediction error than Ordinary Kriging in Sicily as part of an initial study. The focus of the investigation was to use the UK model to interpolate the residuals of a Neural Net Model (Cellura et al., 2008). Qian et al., (2014) utilized UK successfully for wind speed interpolation and prediction in a Bayesian model.

This thesis combines UK and Kalman Filter in two comparative post-processing schemes. The difference in calculation of Universal Kriging is listed below. The basic objective of Kriging is to find the Best Linear Unbiased Estimate (BLUE) of a random field $y(x)$ with n different observations in a field. Kriging interpolation $\widehat{m}$ at location x can be calculated using BLUE (Anagnostou 2015):

$$\widehat{m}(x) = \sum_{i=1}^{n} \lambda_i y(x_i) + \mu(x) \qquad (2)$$

Where λ_i, $i = 1, \ldots n$ are weights obtained from statistical properties that the estimator must possess: unbiasedness and minimum variance (Anagnostou 2015). Different forms of kriging require different first and second moment assumptions based on the equation above.

The second term in Eq. 2 is included to demonstrate the difference in first moment property assumptions between kriging methodologies represented by the mean

field $\mu(x)$. From the above description for Simple Kriging $\mu(x)$= 0, and Ordinary Kriging (and coKriging) require $\mu(x)$to be constant. Universal Kriging allows for trends in $\mu(x)$. The UK mean field is determined by a regression calculated with p chosen functions [$\mu(x) = \sum_{i=1}^{p} \beta_i f(x)_i$] (DACE manual; Coakley et al., 2008). In this study, coefficients β_i, $i = 1 \ldots p$ are regression parameters corresponding to the second order polynomial terms f. Constant and first order regressions were also tested but the results of the In-Sample (IS) application were inferior from the second order regression. Note that p of the terms have been used to create the regression functions in the equation above, leaving p-n terms to describe the local behavior (Cellura et al., 2008; Mardia et al., 1998).

Data comes as a point estimate of hourly wind or bias. Regression and spatial correlation information is calculated hourly. Regression is second order anisotropic, and the semivariogram is calculated with an exponential model. Spherical models are popular with wind speed due to their reliance on the nugget effect (Luo et al., 2008). However exponential models were used to combine spatio-temporal information in (Qian et al., 2014) with a high degree of accuracy. To provide the maximum amount of wind information, all hourly data is used for Kriging.

Stations with gaps of missing data larger than 5 hours within a storm (2.1 Observations) were omitted from the accuracy calculations. Due to the scale of the study, the Kriging model cannot differentiate between very close stations. If two stations were less than 1 km apart their time series for that storm are averaged. This allows for the preservation of original in situ wind information as much as permitted by the model. Note the gridded Kriged output matches the WRF resolution of 2 km.

The Universal Kriging model used in this study is DACE MATLAB Kriging Toolbox (Version 2.0, August 1, 2002 Hans Bruun Nielsen (hbn@imm.dtu.dk) Technical University of Denmark DK-2800 Kgs. Lyngby – Denmark). Other research studies using this model can be seen in (Ryu et al., 2002) and (Coakley et al., 2008).

3.2 Systematic Error Removal (Kalman Filter)

Wind speed prediction bias is reduced using the Kalman Filter. An in depth description of the procedure can be found at (Welch Bishop 2006). Kalman Filters have proven to be effective tools for bias reduction of NWP wind speed predictions (Louka et al., 2008; Delle Monache et al., 2011; Galanis et al., 2006). Correlated bias is minimized with recursively updated weights generated from forecasts and recent observations (Galanis et al., 2009). The Kalman Filter uses a short series of background information and is capable of rapid adjustment (Galanis et al., 2009). This is particularly relevant to the bias of predicted wind speed, which is affected by temporal factors such as season and the diurnal cycle (Wyszogrodzki et al., 2013).

A unified description of the Kalman Filter as a predictor corrector for wind speed is given by (Ide et al., 1997) and used in (Galanis et al., 2009; Galanis et al., 2006; Emmanouil et al., 2012). The Kalman Filter estimates the state of a discrete time controlled process whose true value at time t_i is denoted here by x^t (Eq. 3). This state vector is related to known (WRF bias) observations y_i^o by Eq. 4. This relationship is commonly referred to as the observation equation.

$$x^t(t_i) = M_{i-1}[x^t(t_{i-1})] + \eta(t_{i-1}) \qquad (3)$$

$$y_i^O = H_i[x^t(t_i)] + \varepsilon_i \qquad (4)$$

The matrices M (system operator) and H (observation operator) have to be determined before the application of the filter (Galanis et al., 2006). The random variables η and ε are state and measurement noise. They are assumed independent, white and normal (Welch Bishop 2006). The state noise covariance Q and measurement noise covariance R are calculated with each new set of data. Measurement noise covariance R provides a means of weighting the need for the model innovate with the newest data (Eq. 9) and state noise covariance provides a means of updating the forecasted state error covariance (Eq. 6).

First a forecast step of the state vector $x^f(t_i)$ and its error covariance matrix $Q(t_i)$ based only on the previous time step analysis values are generated by Equations 5 and 6:

$$x^f(t_i) = M[x^a(t_{i-1})] \qquad (5)$$

$$P^f(t_i) = M_{i-1}P^a(t_{i-1})M_{i-1}^T + Q(t_{i-1}) \qquad (6)$$

This forecast is analyzed by incorporating observations available at time t_i. With this information, the state update (Eq. 7), state covariance update (Eq. 8) and Kalman gain (Eq. 9) equations are shown below.

The Kalman gain (K_i) determines the effect the previous bias has on the forecast. Notice that as the measurement error covariance R approaches zero the gain weights the residual more heavily. As the state estimate error covariance P approaches zero the

gain weights the residual less heavily (Welch Bishop 2006). In this notation, *o* denotes an observed value, *t* denotes a true value, *a* denotes analysis and *f* denotes forecast.

$$x^a(t_i) = x^f(t_i) + K_i\left(y_i^O - H_i[x^f(t_i)]\right) \quad (7)$$

$$P^a(t_i) = (I - K_i H_i)P^f(t_i) \quad (8)$$

$$K_i = P^f(t_i)H_i^T[H_i P^f(t_i)H_i^T + R_i]^{-1} \quad (9)$$

There are several different types of the Kalman Filter. A few are discussed in (Welch bishop 2006). The approach used in this thesis is a discrete linear filter. The Kalman Filtered prediction should be statistically more accurate in the least square sense than the WRF raw forecast at the corresponding time (Delle Monache et al., 2006). The Kalman Filter is run for each hour of the day in order to respect the diurnal cycle. The state vector x^t and observation vector y^o contain values corresponding to an hour from 00 to 23 UTC.

Since the bias correction is additive, the Kalman Filter predicted corrector was given bounds in order to avoid negative forecast values similar to (Delle Monache et al., 2006). The Kalman Filter is used in this study through 108 storm events producing 107 filtered storm outputs due to the first storm being as a training period. Kalman Filter iterations described in the next section of this thesis will be referred to as time steps. Arrays of 00 to 23 UTC information will be described as "*sets*". The difference between the WRF raw forecast and Kalman Filter will be referred to as the "expected residual" in the next section.

3.3 Combining Universal Kriging (UK) and Kalman Filter (KF)

3.3.1 The First Methodology

For this thesis two methodologies are compared in order to discuss their effect on bias reduction. The first methodology interpolates all available hourly METAR and MADIS in situ wind speeds to a gridded domain that matches the WRF forecast domain. The KF is then run through each grid cell independently. Figure 3 illustrates this process in terms of KF time steps. The observed wind speeds "O" (OBS) are interpolated to produce a Kriged wind speed "m" in (KGOBS) for hours 1-24 of the history denoted T1 (the first "training" time step). The differences between "m" and "w" forecasted wind speed (WRF) for a given location become y_i^0 in the KF observation equation (Eq. 4). The filter produces a set of expected residuals which are then added to the wind speed forecast the next day WRF(T2), to remove the systematic bias. The corrected WRF forecast of wind speed is denoted "K" in KFWRF(T2).

State information x^t used to calculate the expected residual H_i $[x^a(t_i)]$ at T2 is then updated with the measured bias information at T2. This update estimates the T3 expected residual which is added to WRF(T3) for corrected forecast to produce "K" in KFWRF(T3). The Kalman Filter repeats the forecasts and analysis for the rest of the history, incorporating more information as the time series progress. Method 1 assumes that the Kriged values are accurate. All interpolated observations "m" are treated as accurate as observations "O". The Kriging model simply provides a method of creating a history of observation-model pairs at every location.

Figure 3: Schematic depiction of Methodology 1.

3.3.2 The Second Methodology

Rather than interpolate the observed wind speeds, the second methodology uses Universal Kriging to interpolate the expected residuals. The Kalman Filter observation vector y^0 is composed of the difference between WRF and in situ OBS for time 1-24 (the first time step T1). This bias behavior is used to produce the expected residual output (KFWRF "B") to the corresponding location at the next time step (T2). Notice the expected residuals are only available at the locations there were observation data at the previous time step. Expected Residuals are then interpolated to every grid cell center in the WRF forecast domain, hourly (E). The Kriged expected residuals are still bias not wind speed. The WRF forecast at the associated time step is then added back to the interpolated expected residual in order to produce an adjusted forecast referred to as Residual Kriged Wind (R) at all grid cells.

Temporal updating of the Kalman algorithm occurs every time step for the entire station-WRF pair history for each location. The continuation of updating and producing filtered values for the whole history is described in terms of the update from T2 to T3 as follows. State information used to calculate the expected residual at T2 is then updated with measured bias information at T2. The update estimates the T3 expected residual to produce B(T3), which then is added back to WRF forecast to produce T3 Residual Kriged wind speeds R(T3). The process is not illustrated in the figure below because the forecast and update through a time series was discussed in the previous section.

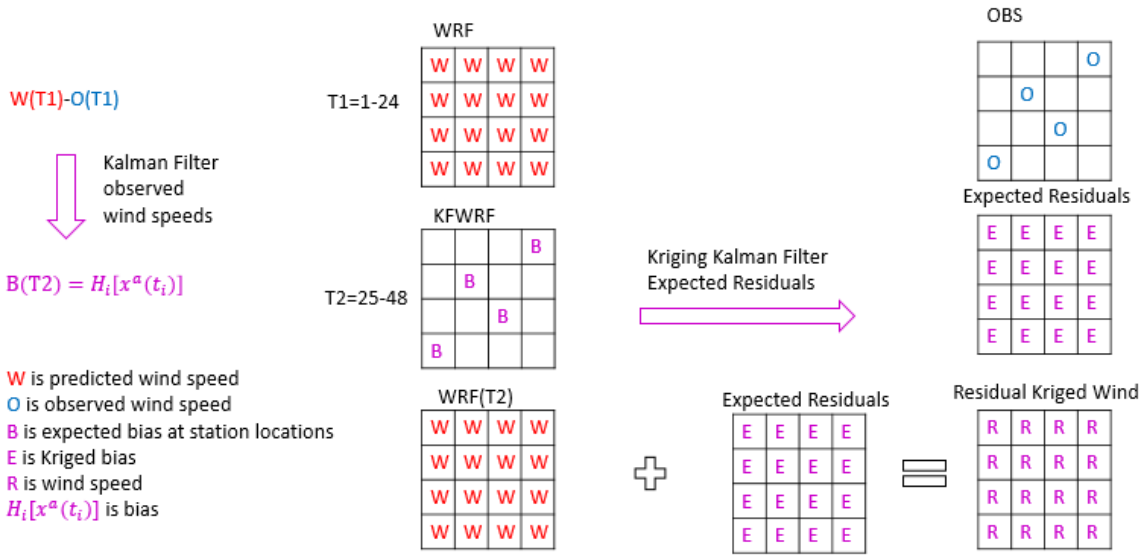


Figure 4: Schematic depiction of Methodology 2.

By interpolating expected residuals with UK, the final correction to the WRF raw forecast undergoes a rudimentary combined temporal and spatial analysis. By using expected residuals as inputs to Eq. 2, the UK model analyzes the second order anisotropic regression and spatial correlation behavior of the expected residuals

hourly. This is similar to a combined methodology (Al-Awadhi and Ali 2012; Cortes 2009) known as the Kriged Kalman Filter (KKF). The difference between Method 2 and KKF is KKF Kalman coefficients are found with respect to the maximum likelihood given spatial covariogram parameters (following one possible procedure and after a number of assumptions and detailed analysis to produce the spatial fields). In this way spatial and temporal covariance are considered simultaneously for the KKF. In contrast, the second methodology of this thesis is an interpolation of all available corrections. It can be broken into two steps, a temporal analysis, and then a spatial analysis of the temporal outputs.

The second methodology differs in the assumptions introduced to the KF compared to the first methodology. In the first methodology, interpolated wind speed error is introduced before the filter has begun. In the second methodology the original observations are used. Also after the expected residuals have been produced all hourly corrections are analyzed for trends and correlation that can incorporate past bias information from surrounding expected residuals.

3.4 Data Processing and Accuracy

In order to capture the difference in performance between the 107 storms and various stations used in each storm, station bias (Eq. 10) and RMSE (Eq. 11) are calculated (using each hour of the storm as i). Bias and RMSE were calculated for UK for wind speed, UK for existing residuals, In-Sample UK wind speed interpolation, Residual Kriging, Methodology 2 Kalman Filter and Methodology 1 Kalman Filter. Individual model output for evaluation is denoted MO

in the equations below. Station storm averages were also used in scatter plots. Eq. 12 and Eq. 13 generate points evaluated for best fit line. This analysis covers a variety of evaluations. The RMSE is generally more sensitive to large errors because the squared difference provides more weight to larger errors. The Bias gives a difference in the central location (Wyszogrodzki et al., 2013). Correlation coefficient evaluates the relationship between forecasts and observations but lacks information of the bias and variation of that relationship.

$$Bias = \frac{1}{n}\sum_{i=1}^{n}(MO_i - Obs_i) \quad (10)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(MO_i - Obs_i)^2} \quad (11)$$

$$Obs_{station} = \frac{1}{n}\sum_{i=1}^{n}(Obs_i) \quad (12)$$

$$MO_{station} = \frac{1}{n}\sum_{i=1}^{n}(MO_i) \quad (13)$$

In order to measure the accuracy of UK interpolations at locations without station inputs, a process called Leave one out Cross Validation (LOOCV) was performed. LOOCV is a popular form of evaluation of wind speed interpolations (Luo et al 2008; Joyner et al., 2015; Zlatev et al., 2010). The procedure involves removing each (applicable) station from the hourly UK input independently and performing the Kriging. In order to accommodate the large amount of station hour combinations LOOCV was run on a smaller database of storms. Eleven storms were selected out of the 107 storm database (10%) randomly using a uniform distribution. This is similar to Luo et al., (2008) performing LOOCV over 10 randomly selected days between 1998 and 2002 to compare wind speed interpolation methods.

Certain data considerations were made in order to produce the maximum data availability for Semi variance calculation, all hourly wind speeds or expected residuals were used for interpolation calculations. Accuracy evaluations performed in this investigation require full storm history of observations. Due to missing values in station data, incomplete histories within a station and storm exist. Stations with missing information are used for interpolation regardless of temporal inconsistencies. However LOOCV accuracy statistics was only performed with the stations without temporal inconsistencies.

LOOCV was performed to evaluate Method 1 UK interpolations of wind speed (UK CV). The interpolation-observation pairs created at every available location for the 11 randomly selected storm events provides a means of calculating overall accuracy statistics and creating a scatterplot representation of Method 1 UK accuracy. The same set of interpolation-observation pairs was used to represent the geographic distribution of wind speed interpolation errors. This allows for the identification of troublesome areas affected by geographic properties or data availability.

Further evaluation of Method 1 interpolation accuracy spanned comparing LOOCV interpolation-observation pairs to another set of interpolation observation pairs calculated with knowledge of the wind speed at the location. This process is known as In–Sample Kriging (UK IS). UK IS in this study has been used to identify if spatial correlation functions represent wind speeds in a similar way with and without data at the location of evaluation. Comparisons between UK CV and UK IS spanned bulk statistics and scatter plots as well as the geographic distribution of error. UK IS was also analyzed for the impact on the Kalman Filter. The same storms and locations used to evaluate UK IS were used to calculate UK IS, M1 KF and WRF wind speed bias. The three populations of bias were examined in terms of relative likelihood of

bias, expressed as a probability density function (PDF). Further population behavior was examined through quartile behavior, represented with boxplots.

Method 2 interpolation was evaluated with a LOOCV of the KF residuals ("B" in Fig. 4) compared to a blind interpolated representation ("E" in Fig. 4), referred to as Leave one out Cross Validation of residual Kriging (RK CV). RK CV was performed on the same 10% of storms randomly selected in the Method 1 LOOCV (UK CV). Note there is a slightly reduced population of KF-RK CV pairs compared to the number of observation-UK CV pairs due to the difference in time step between the expected residual output and in situ wind available for comparison. This is explained in more detail in the results section.

The KF residuals and their interpolated representations were evaluated with bulk statistics and scatter plot as residuals. Further investigation of the effects of interpolating residuals spanned the bias of the interpolated residual as a wind speed prediction (M2 CV). The interpolated residual was represented as a wind speed by adding the associated WRF forecast. KF expected residuals (B in Fig. 4) were also added to the associated WRF forecast (M2 KF). The effect of bias reduction from original WRF forecast, before (M2 KF) and after (M2 CV) interpolation were compared using the PDFs and boxplots of bias and RMSE of the three types of wind speed calculation.

After the effects of interpolation on each methodology was performed, the remainder of accuracy measurements performed in this study was done with all 107 storms. The two post-processing methodologies were compared to each other and the raw WRF forecasts with bulk statistics and scatterplots of predicted-observed wind speed pairs. Further evaluation spanned the geographic distribution of bias and RMSE of each prediction to identify areas of improved

performance. The effect of interpolation on bias and RMSE population was evaluated with the PDFs and boxplots of each prediction method.

4. Results

4.1 UK Wind Speed Interpolation (Method 1)

UK performance is essential to the functionality of both methodologies. The first methodology interpolates observed wind speed to every grid cell. UK CV and UK IS accuracy measurements indicated oversimplification of the wind speed interpolations. Spatial analysis of the error and bias of these results indicate low data availability areas such as the Massachusetts coast and Long Island to be a contributing factor in high bias and RMSE. The effect of oversimplification on bias reduction is evident in the bias population comparison of the M1 KF, UK IS and WRF. Bias populations are compared with PDFs and boxplots. The M1 KF bias is reduced from WRF but with similar nonzero peaks seen in the UK IS PDF. All behavior measured in this section is of 11 of the 107 storms (10%).

Table 1: Accuracy Statistics of UK interpolations of Wind Speed

	UK CV	**UK IS**
RMSE (m/s)	1.34	1.16
Mean Bias (m/s)	-0.66	-0.16
Corr. Coeff. R	0.83	0.85
Slope of linear regression	0.53	0.58

(a) (b)

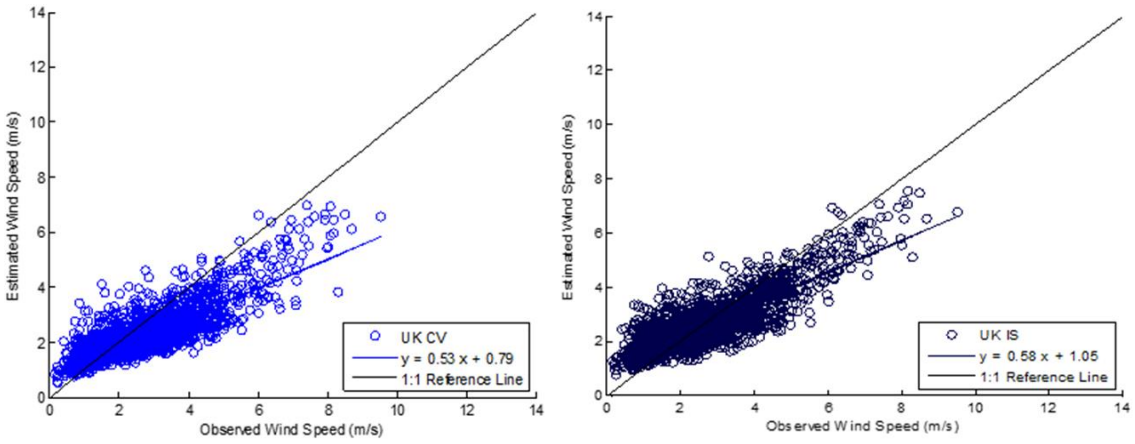


Figure 5: Scatter Plots of (a) UK Leave one out Cross Validation Wind Speed and (b) In-Sample UK Wind Speed vs. Observed Wind Speed

Leave one out Cross Validation results of the first methodology (UK CV) are plotted in Figure 5a. Each point shows the storm average blind interpolated wind speed plotted against the observed storm average wind speed giving each stations performance of each storm. UK

interpolated wind speeds captured observed wind speed behavior with notable consistency (correlation coefficient=0.83). However, a reduction in representation of wind speed variability is reflected in slope of the best fit line (0.53), RMSE (1.34 m/s) and Bias (-0.66 m/s). This could be due to a number of factors, data availability being the largest difference between this and more successful studies.

In-Sample UK (UK IS) results are shown in Figure 5b. The same locations and storms are interpolated with knowledge of the observed wind speed. Similar behavior between UK CV and UK IS can be seen in Table 1. The In-sample correlation coefficient was 0.85 and there was a similar pattern of simplification (RMSE=1.16 m/s) and slope of the best fit line=0.58). The lower magnitude bias (-.16 m/s) is expected as a result of the unbiased requirement of kriging. The existence of the bias is due to the differences between stations and grid cell center locations.

The geographic distribution of UK CV RMSE (Fig 6a) and UK IS RMSE (Fig 6b) show the spatial distribution of wind speed interpolation error. Larger errors in the wind speed interpolations are grouped spatially into the same locations for both UK CV and UK IS. The properties of both wind speed interpolations show high error where there is limited data such as the Massachusetts coast and Long Island. This indicates much of the kriging error represented in the statistics is coming from these areas. The similar results for UK IS and UK CV indicate even with an observation close by there is still a dependence on surrounding information not available due to the lack of oceanic data for both UK CV and UK IS.

(a) (b)

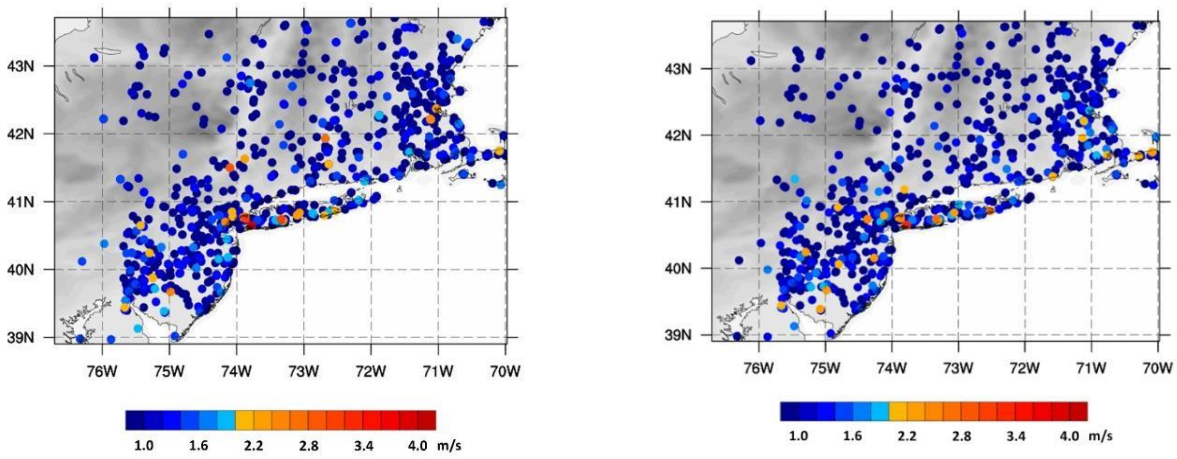


Figure 6: Geographic distribution of RMSE from (a) LOOCV UK (b) In-Sample Wind

Simplification of wind speed behavior varies spatially as shown in the geographic distribution of RMSE for both UK CV and UK IS. Wind speed interpolation accuracy is essential to the first methodology because interpolated wind speeds are assumed to be as accurate as observations. The M1 Kalman Filter removes systematic biases are between WRF and Kriged observations ("m" in Fig. 3) with no consideration to the uncertainty of the interpolation. The effect this has on the Kalman filtered outputs (M1 KF) in relation to observed wind speed is shown in Figure 7. The bias was taken at every available station, and averaged by storm for the WRF, UK IS, and M1 KF results to produce three populations of bias that are represented by their PDFs (Fig 7a) and boxplots (Fig 7b).

With the inclusion of observation data there should be little to no UK IS bias. The boxplot of wind speed interpolation bias shows the limitations of Kriging wind speed over a domain this size with the data available. The UK IS PDF illustrates the effect of oversimplification of the wind speed when performing interpolation identified in the previous estimations (Fig. 5 and 6). The model generally speaking over or under predicts the wind speed with a greater frequency than a zero bias. When the M1 KF reduces the WRF forecast bias with respect to UK IS the nonzero bias peaks from UK IS are introduced to the M1 KF PDF illustrated with the green probability density function of M1 KF bias. Which is between the higher bias WRF (blue) and UK IS, with the same peaks as the UK IS.

(a) (b)

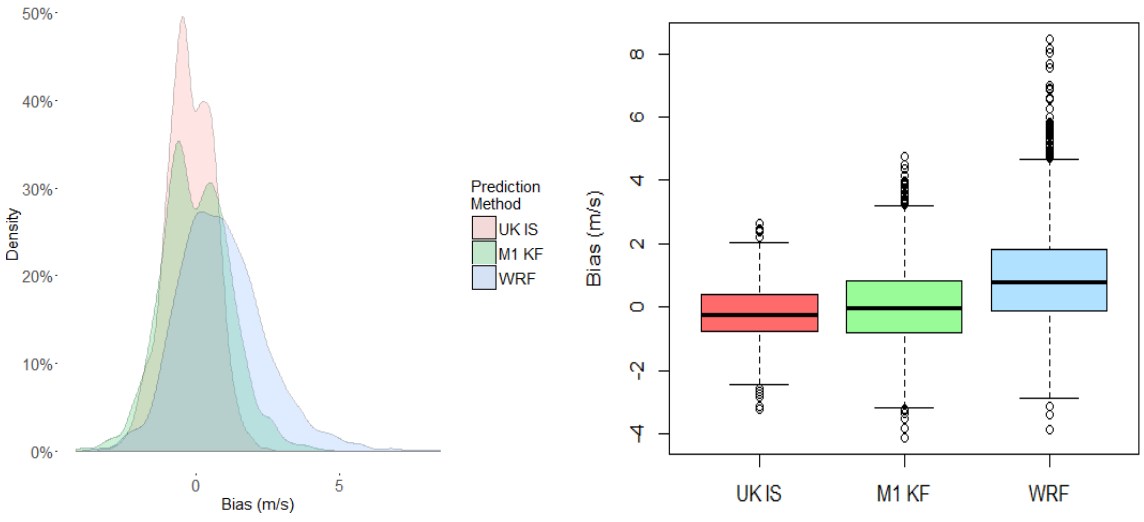


Figure 7: First Methodology (a) Boxplot Bias (b) PDF Bias

4.2 RK Residual Interpolation (Method 2)

Universal Kriging interpolates expected residuals to provide the final predictor correction added to the WRF forecast in the second methodology. RK CV results indicate a simplified but highly correlated interpolation of expected residuals originally created by the Kalman Filter. The effect of the expected residual simplification on a reconstructed wind speed (M2 CV) is a wider PDF of the bias but a better median and lower RMSE compared to a KF reconstructed wind speed before interpolation (M2 KF). All results shown in this section are of the 10% sample of the total storm history used to evaluate Method 1.

RK CV and UK CV studies differ in the independent variable and station-storm pool available for the 11 storm analysis. There is a time lag of 24 hours between observation-WRF pairs used to train the Kalman Filter and the expected residuals Kriged (RK CV). Due to the need for observations at the time step analyzed in CV and the time step before, there is a slightly reduced station pool available for RK CV. The results of RK CV are shown in Figure 8 with associated accuracy calculations in Table 2. Each point in Figure 8 shows the average blind interpolated wind speed plotted against the average expected residual produced by the Kalman Filter. In this study UK was very consistent in predicting the overall Kalman behavior with a correlation coefficient of 0.9. The RMSE (1.01 m/s), bias (-0.11 m/s) and slope of the best fit line (0.56) indicate the expected residuals are simplified in the process of interpolation. Similar to the first methodology, there is a reduction in the ability to capture the variability of the expected residuals.

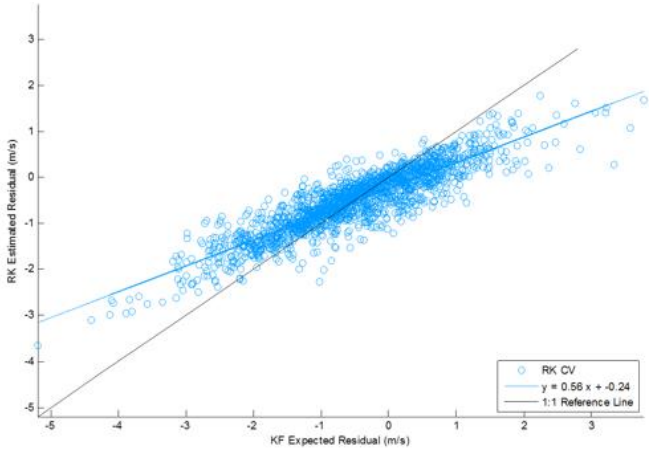


Figure 8: Leave one out Cross Validation of Expected Residuals

Table 2: Accuracy Statistics of UK interpolations of Expected Residuals

	RK CV
RMSE (m/s)	1.01
Mean Bias (m/s)	-0.11
Corr. Coeff. R	0.90
Slope of linear regression	0.56

Results of RK CV indicate a consistently similar but simplified representation of expected residuals through interpolation. The impact of RK simplification on Method 2

accuracy is evaluated by comparing CV corrections to their associated wind speed. This is done by adding the RK CV corrections to associated WRF forecasts (M2 CV). In order to get a baseline performance before interpolation, the expected residuals produced by the Kalman Filter are added to their associated forecasts (M2 KF). M2 CV, M2 KF and the original WRF predicted wind speeds are compared by associated bias and RMSE populations.

The bias population of M2 KF, M2 CV and WRF are represented by boxplots (Fig 9a) and PDFs (Fig 9b). An initial comparison of WRF and M2 KF show a reduction in the positive bias exhibited by WRF. M2 CV reduction of variability of the expected residuals has two effects. The simplification of predictor correctors provides less overall reduction of bias on a by station basis. However the simplification does provide benefit to some stations because there is a lower median bias for M2 CV compared to M2 KF. Figure 9 (c and d) show the RMSE of M2 KF, M2 CV and WRF. Notice the reduction in M2 KF RMSE in median, first and third quartile and even outlier behavior for the 11 storms. Although the bias reduction is reduced the incorporation of spatial behavior of the expected residuals reduces hourly errors made by the KF.

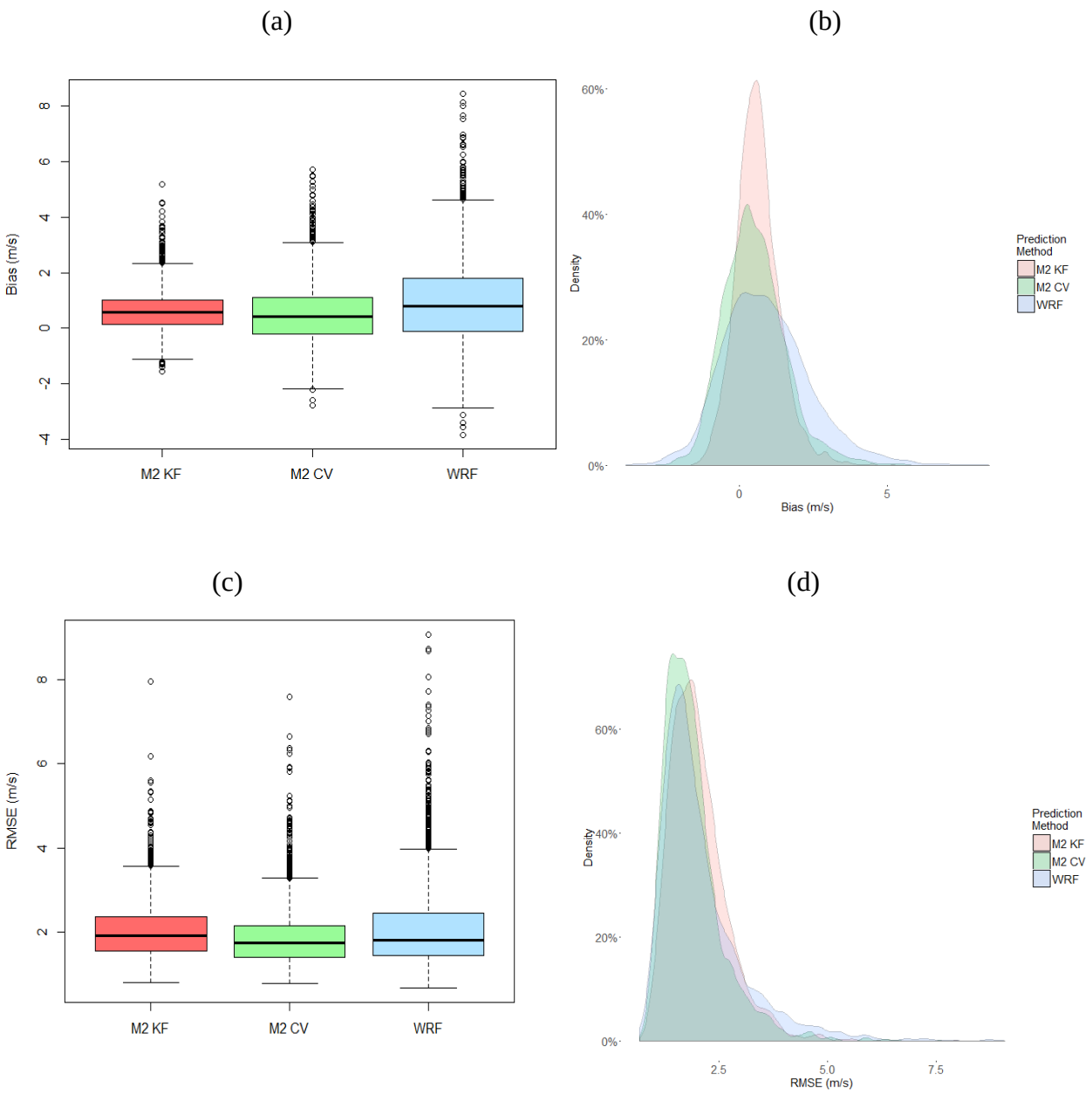


Figure 9: Second Methodology (a) Boxplot Bias (b) PDF Bias (c) Boxplot RMSE (d) PDF RMSE

While UK CV and RK CV results are not directly comparable, both interpolations show evidence of simplification. For the first methodology, interpolations are assumed to be as accurate as observations. Therefore simplification

introduces error. For the second methodology, interpolations represent the underlying second order anisotropic regression and local correlation behavior of all the expected residuals of a given hour. The effect of interpolation on all 107 storms will be discussed in the *Overall Performance* section. In order to fully understand all factors affecting the final model outputs, the factors affecting KF performance will be discussed in the next section.

4.3 Kalman Filter Performance

The Kalman Filter represents WRF bias with the state vector x^t. The vector is composed of 00 to 23 UTC bias state information, updated through a history of sequential storm events. The M2 KF (no interpolated values) with a 24 hour times step size has been shown to reduce the WRF bias 93.88% over 107 storms. Figure 10a shows the average predicted raw WRF forecasted wind speed plotted against the average observed wind speed. Figure 10b is the M2 KF average predicted wind speed plotted against the storm average observed wind speed in order to demonstrate KF improvement without the effects of interpolation. Associated accuracy measurements are in Table 3. The correlation coefficient (0.87), mean bias (0.06 m/s) and slope of the best fit line (0.87) of the M2 KF is improved compared to the original WRF forecast.

(a) (b)

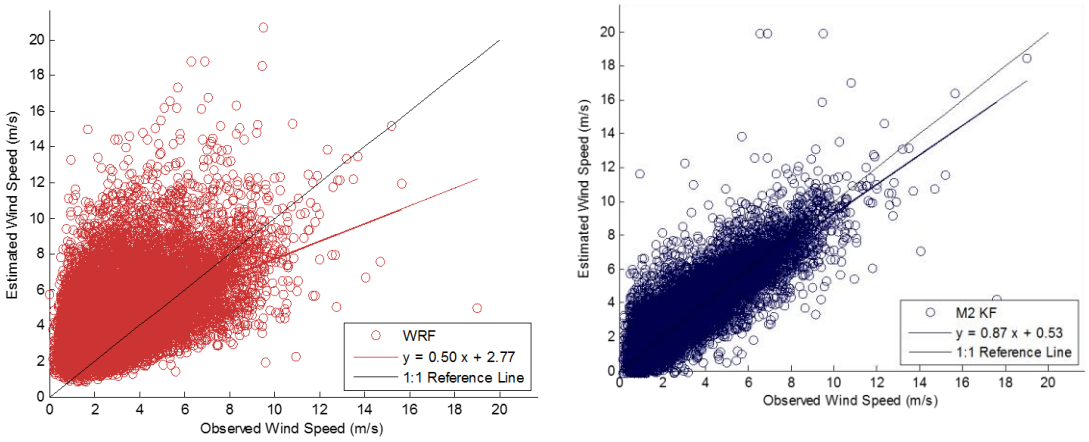


Figure 10: Wind Speed Prediction (a) WRF (b) M2 KF

Table 3: Accuracy Statistics of WRF and M2 KF

	WRF	**M2 KF**
RMSE (m/s)	2.70	2.62
Mean Bias (m/s)	0.98	0.06
Corr. Coeff. R	0.49	0.87
Slope of the linear regression	0.50	0.87

Kalman Filter performance is affected by a number of factors including the time series properties, the time step chosen by the user and effects of interpolation. The Kalman filter represents parameters with initial conditions until the successive updates have reduced their effects. Therefore the performance of the Kalman Filter can be

affected by the number of preceding storms used to train the model. Additionally, the time step of 24 hours captures the diurnal cycle of WRF bias but, improvement in some events would occur with a longer time step used to analyze the bias properties of the storm forecast as a whole. Methodology assumptions affect Kalman Filter training as well. The use of interpolated wind speed (M1 KF) can introduce spatial interpolation inaccuracies into the model when compared to KF outputs without interpolation (M2 KF). However that is not to say that UK wind speed cannot train the Kalman Filter.

Improvement of M2 KF representation of wind speed as the time series progresses is shown in Figure 11. The Kalman Filter is better adapted to reproduce wind speed as the state is recursively updated with more information. Observed wind speed (blue line), WRF predicted wind speeds (red line) and M2 KF wind speeds (green line) are plotted for the first four time steps of station KBOS. The WRF raw forecast has notable systematic under prediction which the Kalman Filter is able to predict more accurately as the time series progresses. The bias of the first day 20031215_00 to 20031215_23 is the difference between the observations and WRF prediction. The second methodology Kalman Filter (M2 KF) is trained for the first day, then the first set of filtered values occur on the second day. The first set of KF outputs removes much of the bias, but a peak during 20031216 _00 UTC creates a spike 24 hours later in the next set of predictions (beginning at 20040702_00). The KF recognizes differences in bias behavior at 00 UTC and decreases the amount of correction applied at 00 UTC. This produces a forecast closer to the observed wind speed at 20040703_00. As the Kalman

Filter progresses through the time series additional information from all previous storms allows for a more refined predicted corrector.

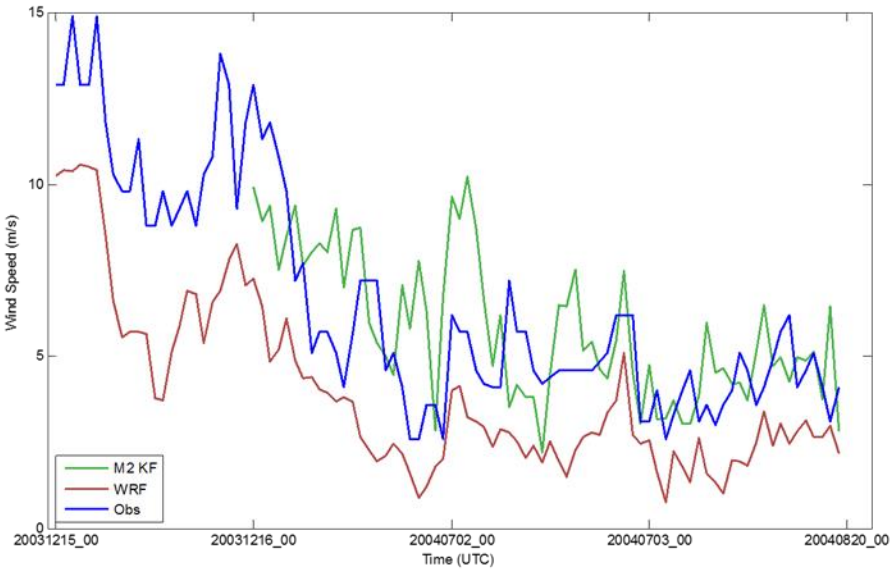


Figure 11: Second Methodology Kalman Filter Training at KBOS

In addition to the amount of training, KF updates are affected by temporal properties of the bias. The time step used in this study is 24 hours in order to quickly adapt to changes in bias behavior and represent the diurnal properties of WRF bias. For certain storms a longer time step could be beneficial. Figure 12 shows four time steps of the station KUKT at a section of data composed of two 48 hour storm events. For both of the storm events there is a peak wind speed during the first day of the model (20040909_00 to 20040909_23 and 20040918_00 to 20040918_23). These the peaks are over predicted by the WRF raw forecasts. This bias behavior is not repeated the second day of the storm event (20040910_00 to 20040910_23 and 20040919_00 to

2040919_23). The different behavior between the bias used for forecasting (at the previous time step) and the actual bias (at the present time step) creates difficulty training the model.

The Kalman Filter cannot capture this 48 hour behavior because it is trained every 24 hours. For the training period 20040909_00 to 20040910_00 there is an under prediction of the peak of the storm for WRF raw forecast. As the storm dies down the WRF raw forecast approaches the observed wind. This presents difficulty for the first Kalman prediction of that time step (green line). There is an anticipated under prediction that is not matched by the observed bias. The KF parameters are then adjusted to reflect the most recent update. Kalman Filtered wind speed values stay very close to the WRF values and miss the similar pattern of WRF over prediction during the peak of the next storm. The 48 hour behavior could be better represented with a 48 hour Kalman filter training. However, note the WRF forecast was relatively close to the wind speed to begin with and the Kalman filter has only undergone a few updates.

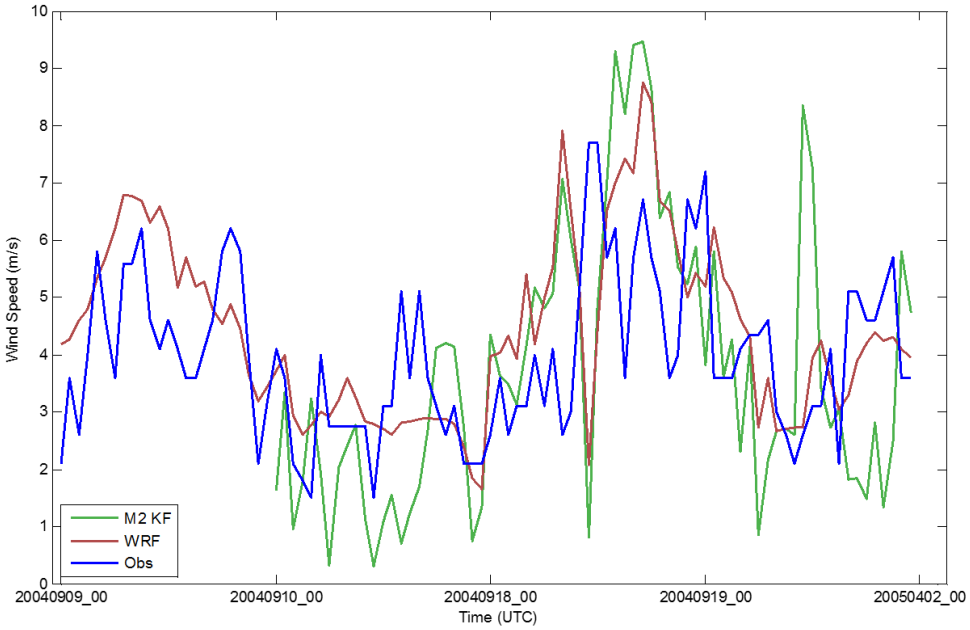


Figure 12: Second Methodology Kalman Filter Training at KUKT

The performance of the Kalman Filter depends on post-processing methodology assumptions (use of Kriging) in addition to temporal properties of WRF bias. Using a UK interpolation as part of the observation vector y^o in the first methodology makes the KF sensitive to the accuracy of UK. Generally speaking UK interpolated wind speed can consistently capture the general behavior of the region; however, the surface cannot fully represent the full variability of the observed wind. This has been evaluated in a cross-validation (Figure 5a) and similar results with In-Sample Kriging (Figure 5b) indicate oceanic data could improve UK accuracy.

A station with a simplified interpolated wind speed used as input to the (M1) KF is shown in Figure 13. The same station (KBOS) data used to evaluate M2 are plotted with the addition of UK interpolated wind speed (black line) and Methodology 1 KF

outputs (purple line). The UK interpolated wind speed throughout both storms is not an improved representation of observed wind speed (blue line). Therefore the bias used to train the M1 KF (difference between the red and black lines) does not provide useful information about the real bias (difference between the red and blue lines). After being trained with an inappropriate representation of the bias for the first day (20031215_00 to 20031215_23) the first predictions from (20031216_00 to 20031216_23) do not bring the M1 KF closer to the observations. This happens with each update of the KF because there is no incorporation of UK inaccuracies into the M1 KF.

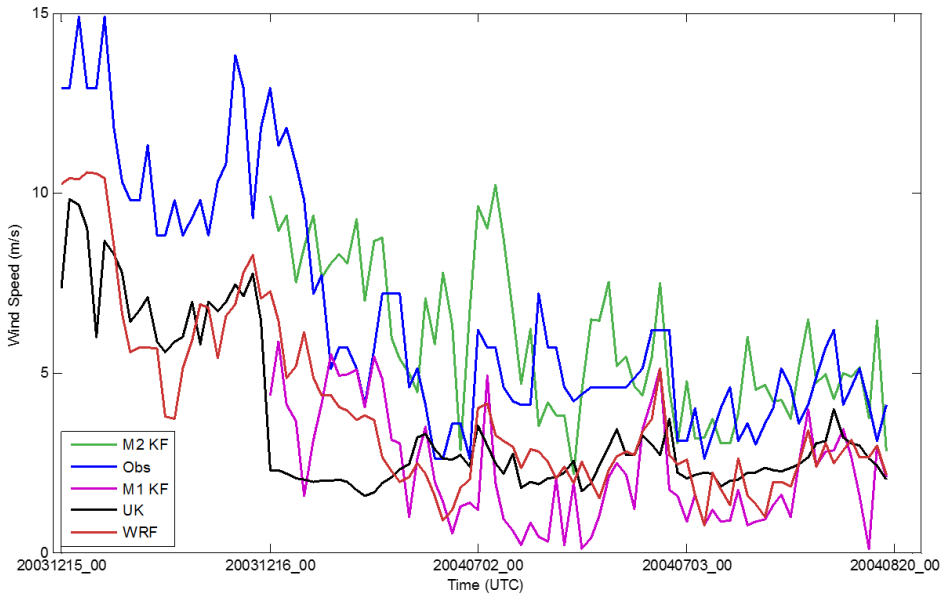


Figure 13: First and Second Methodology Kalman Filter Training at KBOS with UK wind estimates

Accuracy of UK and the impact on M1 KF varied by station and storm. Overall accuracy of both methodologies will be discussed in detail in the next section with final output accuracy statistics in Table 4. Both methodologies were able to reduce the magnitude of the bias. Figure

14 is an example of systematic bias reduction provided by M1 KF. The UK interpolated wind speed over station K3B2 captures the WRF raw forecast over prediction. UK produces a similar wind speed (black line) to the in situ wind speed (blue line). With the high accuracy of UK interpolated wind speed (black line) the Kalman Filtered wind speed (purple line) reduces WRF bias using an interpolated wind speed.

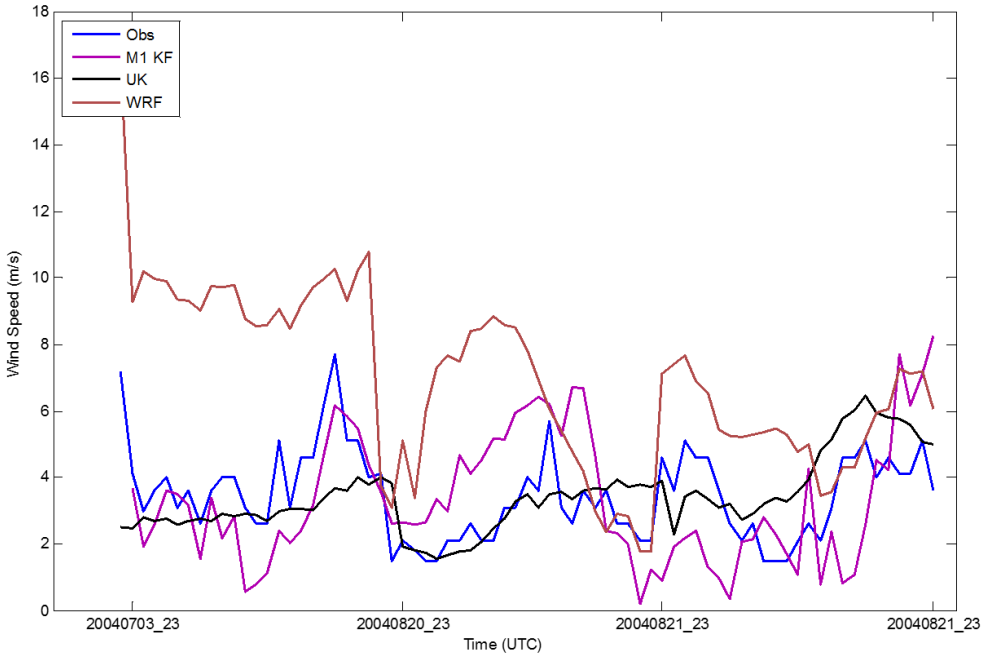


Figure 14: First Methodology Kalman Filter outputs and UK interpolations at K3B2

4.4 Overall Model Performance

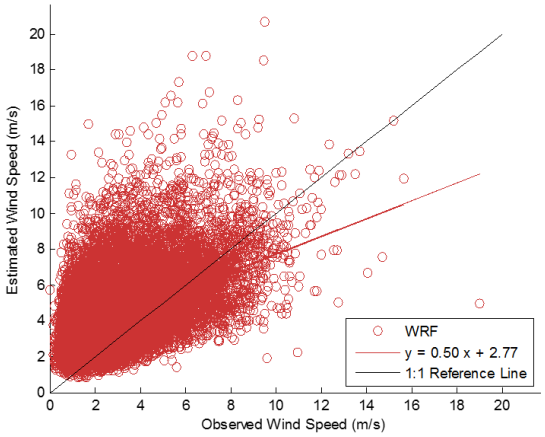
Final post-processing results show regional bias reduction of WRF surface wind speeds using a combined KF and RK methodology. Even with the rudimentary combination of temporal and spatial information, the second Methodology (M2) improved all accuracy measurements when compared to WRF forecasts. The first methodology (M1) reduced raw

WRF forecast bias. However, the use of UK wind speed interpolation reduced ability of post-processed wind speeds to represent the variability of observed wind speed. This dismisses the first methodology as a post-processing feasibility. Limitations in the second methodology include low RMSE reduction and an inability to reduce outlier bias and RMSE behavior. Post-processing performance was not generally effected by season indicating the methodology is capable of reducing a variety of sources of bias.

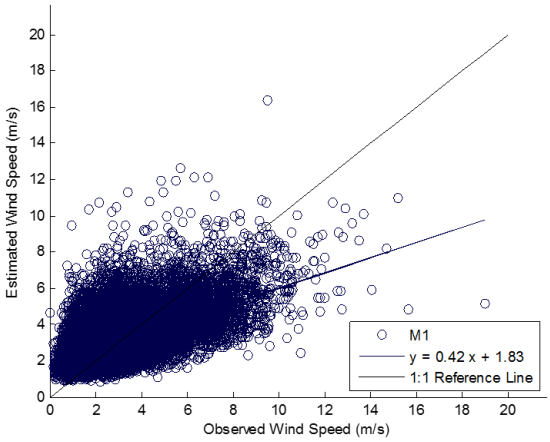
Figure 15 displays storm average estimated wind speed vs. storm average observed wind speed for each station. WRF raw forecasts (Fig 15a), M1 (Fig 15b) and M2 (Fig 15c) are compared to show differences in wind speed prediction. Associated accuracy measurements are shown in Table 4. The effect of UK oversimplification being introduced in to the first methodology Kalman filter can be seen in the reduction of the slope of the best fit line of the M1 output (0.42) compared to the raw WRF forecast (0.50).

The reduced ability of M1 predicted wind speeds to reflect variability of observed wind speed dismisses the methodology as a feasible post processing option. However due to the consistently similar wind speed surface produced by Universal Kriging, an 83.67% a bias reduction did occur. The first methodology results will continue to be discussed in the context of similarities in geographical limitations as well as RMSE and bias properties of a combined Kriging and KF methodology. Method 2 provided improved statistics in all aspects measured in this study. Bias magnitude was lowered 88.78 % without a reduced slope of the best fit line (0.59). This was improved compared to the slope of the WRF best fit line (0.50). There was also an increase in correlation from WRF (0.49) to M2 (0.73). There was some RMSE reduction but it was not significant. The total RMSE reduction was the same for M1 and M2.

(a)



(b)



(c)

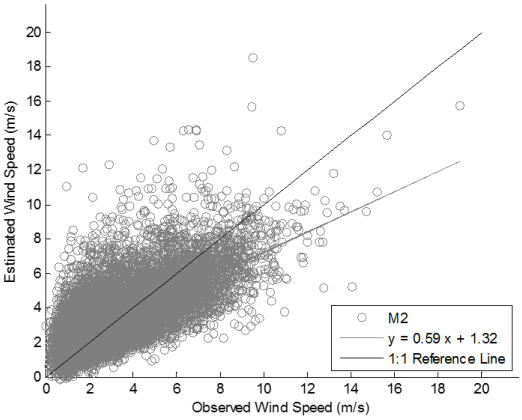


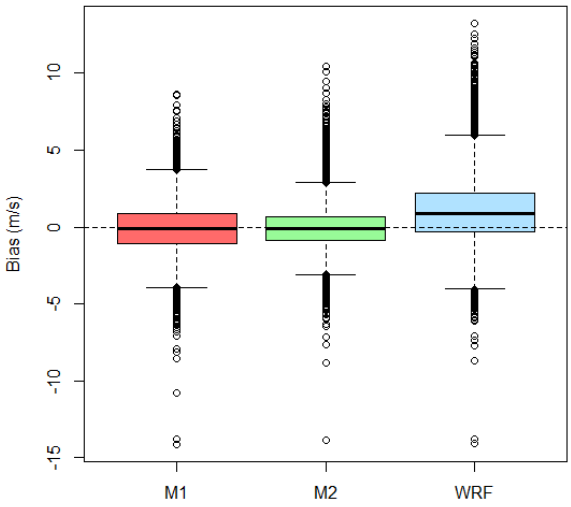
Figure 15: Final Model output (a) WRF forecast, (b) M1 (c) M2

Table 4: Error Statistics of the 107 Storm forecasts

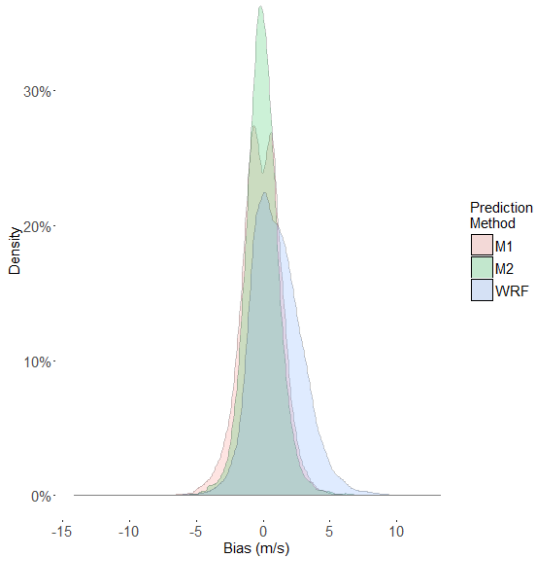
	WRF	**M1**	**M2**
RMSE (m/s)	2.70	2.47	2.47
Mean Bias (m/s)	0.98	-0.16	-0.11
Corr. Coeff. R	0.49	0.61	0.73
Slope of the linear regression	0.50	0.42	0.59

In order to describe the accuracy of the final model outputs in more detail, the PDF and boxplot representation of each methodologies bias and RMSE are shown in Figure 16. Notice for both methodologies the positive bias of WRF was removed (Fig. 16a). This is represented in the reduction of median bias from WRF to M1 and M2. While the overall bias of M1 and M2 are similar (-.16 and -.11 accordingly), the PDF shows the effect of using interpolated wind speeds on the final M1 bias. Evidence of nonzero bias peaks can be seen in the final Method 1 Kalman Filter (M1 KF) probability density function (Fig 16b) first identified in the UK wind speed surface (UK) probability density function (Figure 7b). This creates a wider bias distribution in the boxplot representation of M1 bias compared to M2. The first and third quartile as well as the min and max (not including outliers) of M1 bias reach a higher magnitude when compared to the M2 boxplot. Therefore over the course of 107 storms the bias at a given location is more frequently closer to zero for the second methodology.

(a)



(b)



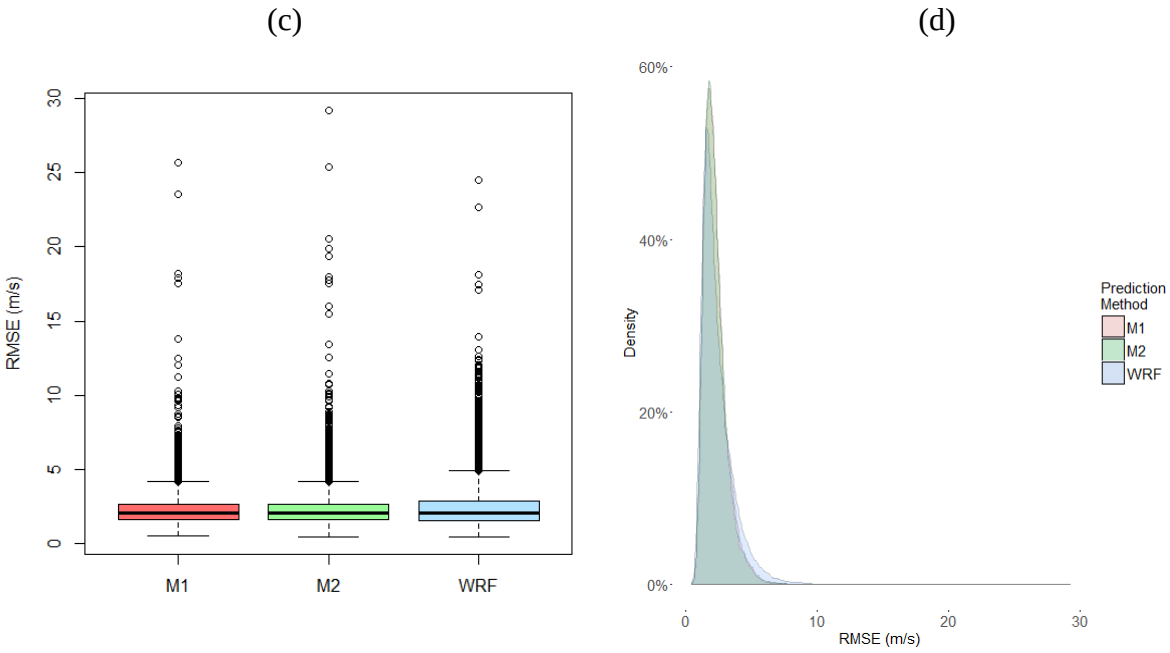


Figure 16: Final output (a) Boxplot Bias (b) PDF Bias (c) Boxplot RMSE (d) PDF RMSE

Limitations of the second methodology include the model ability to capture outliers in the bias and RMSE. This can be seen in the M2 boxplot of the bias. When comparing the M2 bias PDF without outliers (represented by the end of the dashed lines), the distribution has lower positive and negative bias. However outliers in the distribution (points passed the dashed line) reach further for the bias for M2 than M1. This is also true for the RMSE outliers (Fig. 16c). This is due to the inability of interpolation to capture bias behavior dissimilar from its surroundings. This also caused minimal M2 improvement in low RMSE behavior confirmed in RMSE boxplot and PDF (Fig. 16d) compared to M1 and WRF. Method 2 RK captures a large portion of the Kalman Filter expected residual variability but not all. The reason for total RMSE reduction from WRF in Table 4 despite issues reducing outliers, is due to the overall

RMSE reduction in most stations. This was evident in the reduction of the 3rd quartile and maximum (excluding outliers) M2 RMSE behavior.

The limiting factor in the second methodology performance comes from the interpolation. The extent of the expected residual oversimplification is shown in Figure 17. Use of the Kalman Filter without interference from interpolation (M2 KF) yields very significant improvement to raw WRF statistics (Table 5). Notice that the very high correlation (0.87) of the initial Kalman Filter outputs (dark blue points). The success of KF in reducing systematic errors in this study tributes its popularity in other studies (McCollor and Stull 2008; Louka et al., 2008; Delle Monache 2011; Galanis et al., 2006). Similar to the RK CV results, RK is capable of capturing much but not all of this behavior (green points). This can be seen in the less precise behavior of the RK values on the same plot. The added benefit of regional correction is at the expense of overall station bias reduction. A quantitative representation of this effect can be seen with the direct reduction in correlation coefficient and slope of the best fit line between Kalman outputs (M2 KF) and M2 (Table 5).

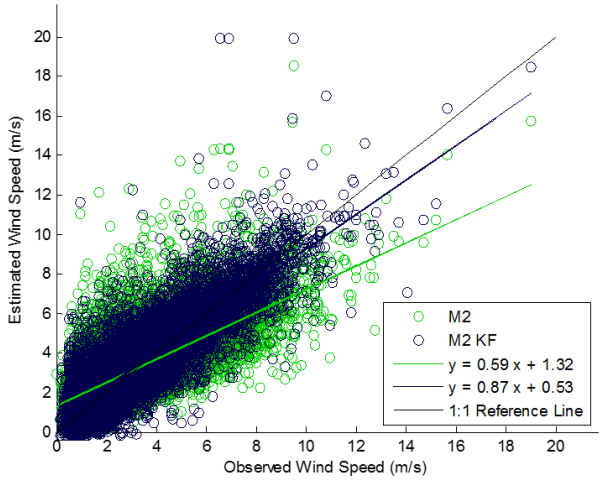


Figure 17: Second Methodology Kalman Filter (M2 KF) and M2 Residual Kriged Wind (M2) comparison

Table 5: Error Statistics of Method 2 KF Before and after interpolation

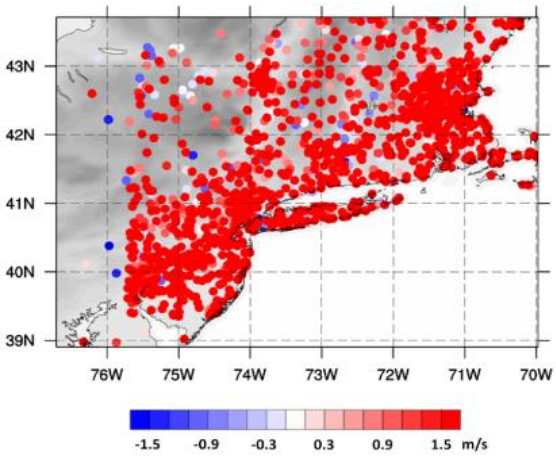
	WRF	**M2**	**M2 KF**
RMSE (m/s)	2.70	2.47	2.62
Mean Bias (m/s)	0.98	-0.11	0.06
Corr. Coeff. R	0.49	0.73	0.87
Slope of the linear regression	0.50	0.59	0.87

The bias spatial distribution of WRF (Fig 18a), M1 (Fig. 18b) and M2 (Fig 18c) is shown below. The WRF raw wind speed forecast before post processing had positive bias over most of the area with emphasis near the coast. There is a notable decrease in the size of areas affected by positive bias for both methodologies of bias reduction. Smaller areas of high bias remained with the second methodology. Although M2 was an improvement

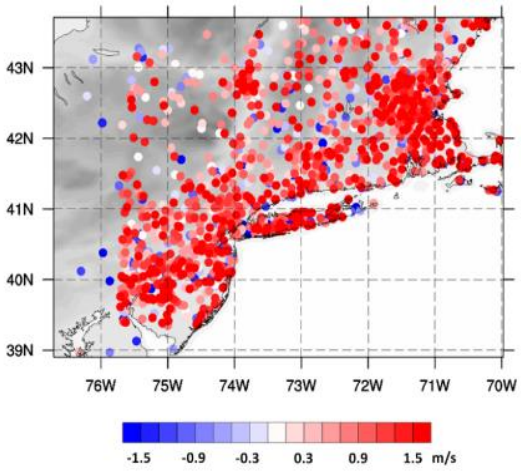
to WRF and M1 there were still less effective areas. High bias remained along the Massachusetts coast and Long Island.

The same areas had persisting RMSE. The geographic distribution of RMSE is shown in Figure 19 for WRF (Fig 19a), M1 (Fig 19b) and M2 (Fig 19c). This is a result of a lack of surrounding oceanic information for these areas. Another noted limitation of the second methodology was a lack of notable improvement in RMSE between M2 and M1. There was minimal difference in M1 and M2 boxplot RMSE behavior (Fig. 16c) and there is similarity of the geographic distribution of RMSE between the first methodology (Fig 19b) and second methodology (Fig 19c). Regardless of whether wind speeds or expected residuals were interpolated the same final RMSE (Table 4) and similar geographic distribution of error was produced.

(a)



(b)



(c)

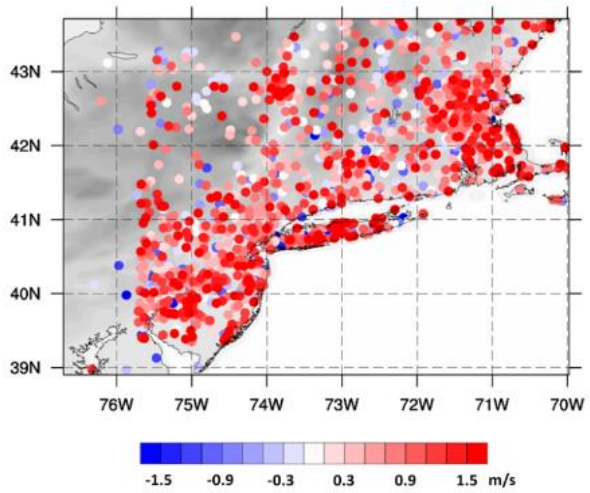
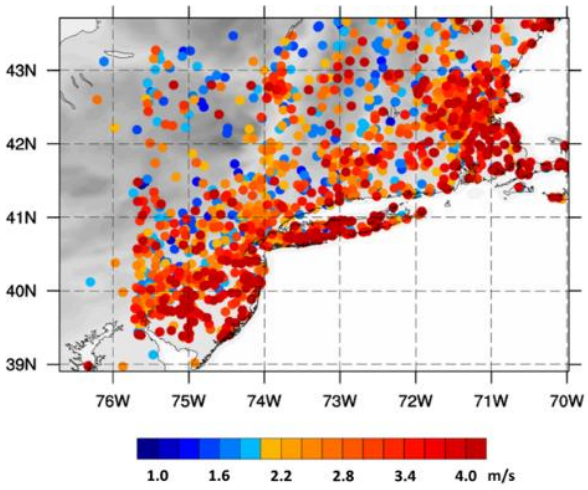
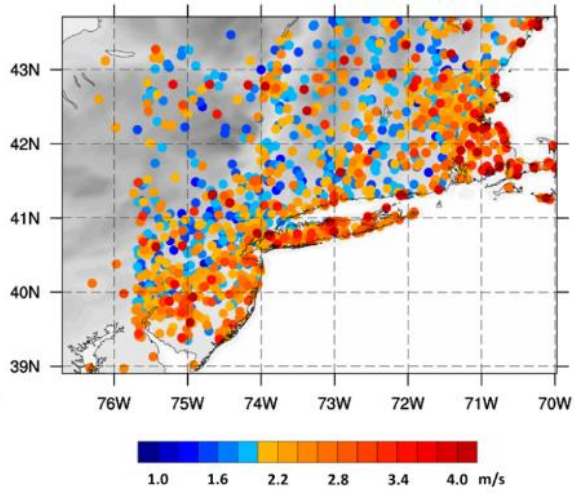


Figure 18: Geographic distribution of Bias from (a) WRF forecast (b) M1 (c) M2

(a)



(b)



(c)

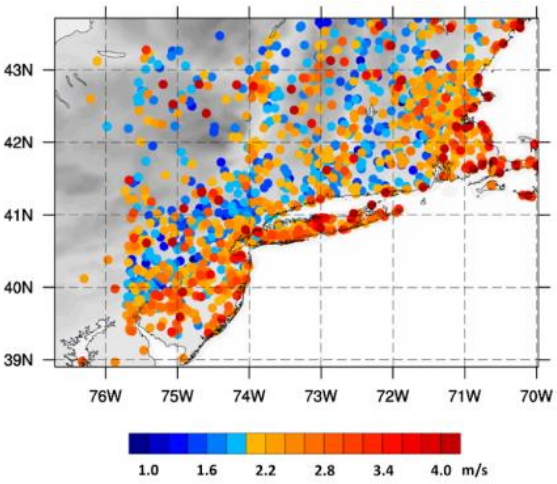


Figure 19: Geographic distribution of RMSE from (a) WRF forecast (b) M1 (c) M2

Both M1 and M2 reduced the overall positive bias of WRF raw wind speed predictions, with improved performance in the second methodology. This was the case regardless of season. The boxplot of WRF, M1 and M2 bias shown in Figure 20a. M2 has a lower magnitude bias for the minimum, maximum, first and third quartiles compared to M1 and WRF. This is the same bias relationship exhibited in the overall final model output bias boxplot (Figure 16a). The consistent bias reduction regardless of season indicates there M2 is able to reduce a wide variety of sources of bias.

Post-processing with the second methodology also had the similar limitations regardless of season. Outliers of the bias distribution were not captured as well for the second methodology as the first methodology (with the exception of a few stations in the summer). The second methodology requires interpolation of expected residuals. Regardless of season certain areas did not exhibit similar behavior to their surroundings. The similar seasonal limitations of the second

methodology indicate a variety of sources can negatively impact M2 performance. Further study to improve these errors will be discussed in the *Conclusions* section. This is due to the KF being a bias reduction tool. Nonsystematic error introduced by the chaotic nature of wind speed was not improved. And the lack of surrounding oceanic data created areas with consistently decreased interpolation performance. The seasonal M2 RMSE (Figure 20b) showed some but not significant reduction in RMSE. This came with the exception of outliers, just as before the results were separated by season.

(a) (b)

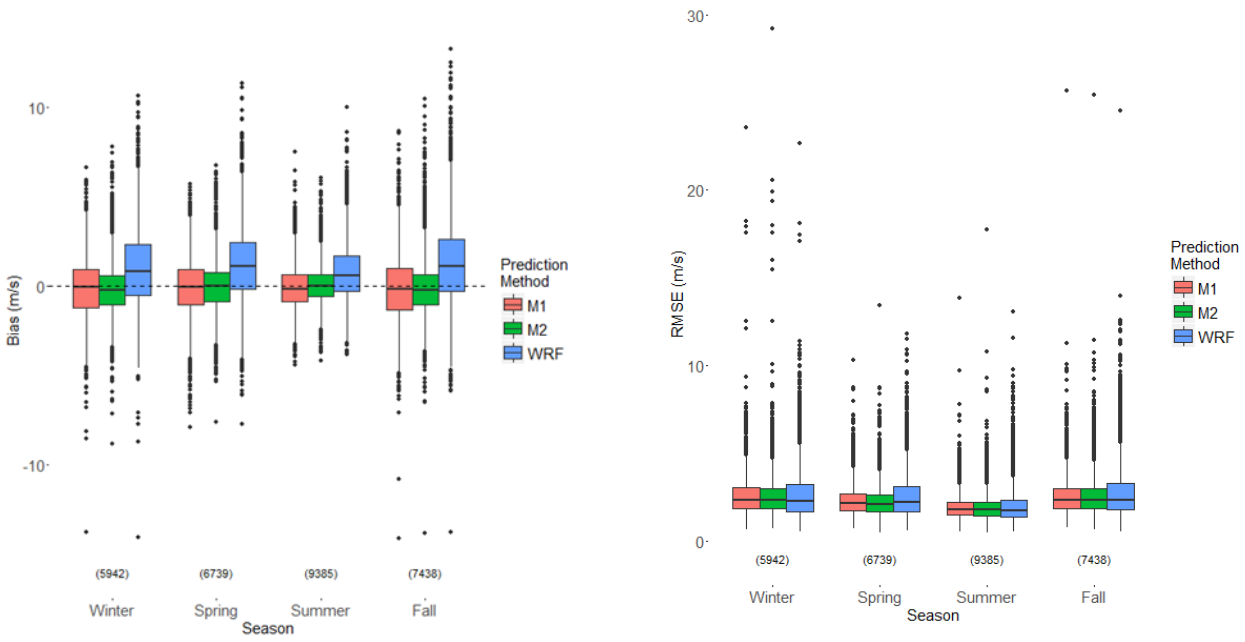


Figure 20: Final Output by Season (a) Boxplot Bias (b) Boxplot RMSE

5. Conclusions

The results of this study indicate regional bias reduction of the WRF forecast can be performed with a post-processing scheme using Universal Kriging to interpolate

the expected residuals produced by the KF over the whole domain (second methodology). This was due to the 93.88% reduction of the bias using a Kalman Filter without the use of any interpolations. Final M2 bias reduction was slightly lower (88.78%) due to expected residual simplification caused by RK. Interpolated expected residuals have been shown to provide improved corrections in some instances; however overall corrections for every location in the forecast domain come at the expense of individual station bias reduction.

Although benefits were not significant enough to suggest implementation on a larger scale, this study has identified the general ability of Universal Kriging to interpolate KF expected residuals for the northeastern United States. Even with the rudimentary combination of temporal and spatial information the second methodology lowered bias, and improved the slope and fit of the best fit line when compared to statistics from WRF raw forecast outputs. Geographically the large area of WRF over-prediction was reduced. Furthermore the same areas with high RMSE geographic distribution were also reduced. Limitations in the methodology come from reduction in Kalman Filter correction variability as a result of interpolation. This effected the second methodology ability to handle outlier behavior in bias and RMSE.

Use of Universal Kriging as a residual interpolator was preferable to a wind speed interpolator (first methodology) due to the impact on final model outputs. Results indicated Universal Kriging to be the limiting factor for the regional bias reduction regardless of methodology. The largest introduction of error occurs close to the coast. This could be reduced if oceanic wind speed data was included. Oceanic data was not included in this study due to the disproportionate error of oceanic data in cross

validations. This was caused by limited coverage of oceanic stations. The inclusion of oceanic data for the purpose of identifying coastal improvement could provide additional insight to current results.

Further interpolation improvement can be made with incorporation of temporal uncertainty in spatial interpolation parameters. Several recent studies have begun to describe the spatio-temporal properties of wind speed (Suryawanshi and Gosh 2015; Tascikaraoglu et al., 2016; Sanandaji et al., 2015; Dowell et al., 2014). Existing spatio-temporal uncertainty calculations exist in the form of the Kriged Kalman Filter, first derived in by Mardia et al., (1998). Spatial and temporal covariance are considered simultaneously for the KKF, as seen in the study performed by Al-Awadhi and Ali (2012) of air quality in Kuwait. This more detailed representation of spatio-temporal residual properties could improve the final predictor correction in the second methodology further.

6. References

1. Akkala, A., V. Devabhaktuni, and A. Kumar. 2010. Interpolation techniques and associated software for environmental data. Environmental Progress & Sustainable Energy 29:134-141.
2. Al-Awadhi, Fahimah A., & Alhajraf, Ali. (2012). Prediction of non-methane hydrocarbons in Kuwait using regression and Bayesian kriged Kalman model.(Report). *Environmental and Ecological Statistics, 19*(3), 393.
3. Anagnostou, Emmanouil N. "Lecture Notes for Probabilistic Methods for Engineering Systems." 2015 Probabilistic Methods for Engineering Systems. University of Connecticut. Lecture.
4. Atkinson and Lee 1992. Procedures for Substituting Values for Missing NWS Meteorological Data for Use in Regulatory Air Quality Models. *Environmental Protection Agency*
5. Cellura, M., G. Cirrincione, A. Marvuglia, and A. Miraoui. 2008. Wind speed spatial estimation for energy planning in Sicily: A neural kriging application. Renewable Energy 33:1251-1266. CGIAR-CSI.
6. Cheng WYY and Steenburgh WJ (2005) Evaluation of surface sensible weather forecasts by the WRF and the eta models over the Western United States. Weather Forecast 20:812–821
7. Coakley, Henry, Joshua Williams, and Doran Baker (2008). "Universal Kriging Interpolator for Satellite Derived Global Data."
8. Cortes, J. (2009). Distributed Kriged Kalman Filter for Spatial Estimation. *Automatic Control, IEEE Transactions on, 54*(12), 2816-2827.
9. Delle Monache, L., Nipen, T., Deng, X., Zhou, Y., & Stull, R. (2006). Ozone ensemble forecasts: 2. A Kalman filter predictor bias correction. *Journal of Geophysical Research: Atmospheres, 111*(D5).
10. Delle Monache, L., Liu, Y., Roux, G., Nipen, T., & Stull, R. (2011). Kalman filter and analog schemes to post process numerical weather predictions. *Monthly Weather Review, 139*(11), 3554-3570.
11. Dowell, Weiss, Hill, & Infield. (2014). Short-term spatio-temporal prediction of wind speed and direction. Wind Energy, 17(12), 1945-1955.
12. Emmanouil, G., Galanis, G., & Kallos, G. (2012). Combination of statistical Kalman filters and data assimilation for improving ocean waves analysis and forecasting.(Report). *Ocean Modelling, 59 60*, 11.
13. Galanis, G., P. Louka, P. Katsafados, I. Pytharoulis, and G. Kallos (2006). "Applications of Kalman filters based on non-linear functions to numerical weather predictions." In *Annales geophysicae*, vol. 24, no. 10, pp. 2451-2460. Copernicus GmbH.
14. Galanis, Emmanouil, Chu, & Kallos. (2009). A new methodology for the extension of the impact of data assimilation on ocean wave prediction. *Ocean Dynamics, 59*(3), 523-535.
15. Grell, G.A. and D. Devenyi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques, Geoph. Res. Let., 29, NO 14., 10.1029/2002GL015311, 2002.
16. Hamill, T. M., 1999: Hypothesis tests for evaluating numerical precipitation forecasts. Wea. Forecasting, 14, 155–167.

17. Heath, Nicholas K., et al. "A simple lightning assimilation technique for improving retrospective WRF simulations." *Journal of Advances in Modeling Earth Systems*.
18. Hong, S. Y., Y. Noh, and J. Dudhia (2006), A new vertical diffusion package with an explicit treatment of entrainment processes, Mon. Weather Rev., 134(9), 2318–2341, doi:10.1175/Mwr3199.1.
19. Jimenez PA, Dudhia J (2012) Improving the representation of resolved and unresolved topographic effects on surface wind in the WRF model. J Appl. Meteor Climatol. 51:300–316
20. Joyner, T., 2013. Optimizing peak gust and maximum sustained wind speed estimates from mid-latitude wave cyclones. Geography&Anthropology.LouisianaStateUniversity,LSUETD,p.169.
21. Joyner, Friedland, Rohli, Treviño, Massarra, & Paulus. (2015). Cross-correlation modeling of European windstorms: A coKriging approach for optimizing surface wind estimates. *Spatial Statistics*, *13*, 62-75.
22. Klawa, M., and U. Ulbrich (2003), A model for the estimation of storm losses and the identification of severe winter storms in Germany, Nat. Hazards Earth Syst. Sci., 3, 725 – 732.
23. Lorenzana, Jesús, et al. "Sensitivity of WRF precipitation field to assimilation sources in northeastern Spain." EGU General Assembly Conference Abstracts. Vol. 17. 2015.
24. Louka, G. Galanis, N. Siebert, G. Kariniotakis, P. Katsafados, I. Pytharoulis, G. Kallos, 2008: Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering, Journal of Wind Engineering and Industrial Aerodynamics, Volume 96, Issue 12, December 2008, Pages 2348-2362, ISSN 0167-6105, http://dx.doi.org/10.1016/j.jweia.2008.03.013.
25. MacEachren AM, Davidson JV. 1987. Sampling and isometric mapping of continuous geographic surfaces. The American Cartographer 14: 299–320.
26. Mardia, K., Goodall, V., Redfern, C., & Alonso, E. (1998). The Kriged Kalman filter. *Test, 7*(2), 217-282.
27. Mass CF, Ovens D (2011) Fixing WRF's high speed wind bias: a new sub grid scale drag parameterization and the role of detailed verification. In: 24th Conference on Weather and Forecasting and 20th
28. Conference on Numerical Weather Prediction, Preprints, 91st American Meteorological Society Annual
29. Meeting, 23–27 Jan, Seattle, WA. Paper 9B.6. http://ams.confex.com/ams/ 91Annual/webprogram/Paper180011.html
30. McCollor, D., R. Stull, 2008: Hydrometeorological accuracy enhancement via post processing of numerical weather forecasts in complex terrain. *Wea. Forecasting*, 23, 131–144.
31. doi: http://dx.doi.org/10.1175/2007WAF2006107.1

32. Michalakes, J., et al. (2005) *The Weather Research and Forecast Model: Software Architecture and*
33. *Performance.* Proceedings of the Eleventh ECMWF Workshop on the Use of High Performance

34. Computing in Meteorology. Eds. W. Zwieflhofer and G. Mozdzynski. World Scientific, 2005, pp 156 – 168.
35. Molders N (2008) Suitability of the weather research and forecasting (WRF) model to predict the June 2005 fire weather for Interior Alaska. Weather Forecast 23:953–973
36. Molders N (2008) Suitability of the weather research and forecasting (WRF) model to predict the June 2005 fire weather for Interior Alaska. Weather Forecast 23:953–973
37. Qian, Hongkun, Tao, (2014) Wind Speed Spatio-temporal Forecasting of Wind Farms Based on Universal Kriging and Bayesian Dynamic Model. POWERCON 2014
38. Roux G, Liu Y, Delle Monache L, Sheu R-S, Warner TT (2009) Verification of high resolution WRF RTFDDA surface forecasts over mountains and plains. In: 10th WRF users' workshop, 20–23 June, 2009, Boulder Ryu, JS., Kim, MS., Cha, KJ. et al. Kriging Interpolation Methods in Geostatistics and DACE Model. KSME International Journal (2002) 16: 619. doi:10.1007/BF03184811
39. Sanandaji, B., Tascikaraoglu, A., Poolla, K., & Varaiya, P. (2015). Low-dimensional Models in Spatio-Temporal Wind Speed Forecasting.
40. Skamarock, William C., et al. "A Description of the Advanced Research WRF Version 3. (2008).
41. Sliz-Szkliniarz, B., Vogt, J., 2011. GIS-based approach for the evaluation of wind energy potential: A case study for the Kujawsko–PomorskieVoivodeship.RenewableSustainableEnergyRev.15, 1696–1707.
42. Suryawanshi, A., & Ghosh, D. (2015). Wind speed prediction using spatio-temporal covariance. Natural Hazards, 75(2), 1435-1449.
43. Tascikaraoglu, A. M., Sanandaji, B., Poolla, K., & Varaiya, P. (2016). Exploiting sparsity of interconnections in spatio-temporal wind speed forecasting using Wavelet Transform. Applied Energy, 165, 735-747.
44. Tewari, F. Chen, W. Wang, J. Dudhia, M.A. LeMone, K. Mitchell, M. Ek, G. Gayno, J. Wegiel, R.H. Cuenca
45. 2004. Implementation and verification of the unified NOAH land surface model in the WRF model. 20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction, pp. 11–15
46. Tieleman, H.W., 1995: "Simulation of Surface Winds for Assessment of Extreme Wind Loads on Roofs," Proceedings of the 9th International Conference on Wind Engineering, ppl 162-1169, New Delhi, India.
47. Welch, Greg, and Gary Bishop. "An Introduction to the Kalman Filter," Department of Computer Science University of North Carolina at Chapel Hill (2006).
48. Wyszogrodzki, A., et al. 2013: Analysis of the surface temperature and wind forecast errors of the NCARAirDat operational CONUS 4-km WRF forecasting system. *Meteorology and Atmospheric Physics*, 122, 125-143, DOI: 10.1007/s00703-013-0281-5.
49. Zlatev, Z., S. E. Middleton, and G. Veres. 2009. Ordinary kriging for on-demand average wind interpolation of in-situ wind sensor data.in EWEC 2009, Marseilles.

50. Zlatev, Z., S. E. Middleton, and G. Veres. 2010. Benchmarking knowledge-assisted kriging for automated spatial interpolation of wind measurements. Pages 1-8 in 2010 13th Conference on Information Fusion (FUSION).